

3-8-00

[Client ref. 56115534-120430]

A  
jc511 U.S. PTO  
09/520677

03/07/00

03/07/00  
U.S. PTOTHE ASSISTANT COMMISSIONER FOR PATENTS  
Washington, D.C. 20231

Sir:

Transmitted herewith for filing is the patent application of Inventor(s) Garry Z. Gu

Entitled: **CREDIT-BASED FLOW CONTROL FOR MULTI-STAGE MULTI-DIMENSIONAL SWITCHING ARCHITECTURE**

Enclosed are:

19 sheets of specification and 6 sheet(s) of drawing(s).  
 An Assignment of the invention to Lucent Technologies, Inc. and an Assignment Recordation cover sheet.  
 A verified statement to establish small entity status under 37 C.F.R. §§ 1.9 and 1.27.  
 An executed declaration/power of attorney.  
 An Information Disclosure Statement and form PTO-1449.  
 Other

The filing fee has been calculated as shown below:

	(Col. 1)	(Col. 2)	SMALL	ENTITY		OTHER THAN A SMALL ENTITY	
FOR:	NO. FILED	NO. EXTRA	RATE	FEE		RATE	FEE
<b>BASIC FEE</b>				\$	<u>OR</u>		\$690
<b>TOTAL CLAIMS</b>	<b>18 -20 =</b>	<b>* 0</b>	<b>x 9</b>	<b>\$</b>	<u>OR</u>	<b>x 18</b>	<b>\$ 0</b>
<b>INDEP CLAIMS</b>	<b>2 -3 =</b>	<b>* 0</b>	<b>x 39</b>	<b>\$</b>	<u>OR</u>	<b>x 78</b>	<b>\$ 0</b>
<b>MULTIPLE DEPENDENT CLAIM PRESENTED</b>			<b>x 130</b>	<b>\$</b>	<u>OR</u>	<b>x 260</b>	<b>\$</b>
* If the difference in Col. 1 is less than zero, enter "0" in Col. 2			<b>TOTAL</b>	<b>\$</b>	<u>OR</u>	<b>TOTAL</b>	<b>\$690</b>

Please charge Deposit Account No. 12-2325 in the amount of \$690.00 to cover the filing fee. The Commissioner is also hereby authorized to charge payment of the following fees associated with this communication or credit any overpayment to Deposit Account No. 12-2325. A duplicate copy of this sheet is enclosed.

Any additional filing fees required under 37 C.F.R. § 1.16.  
 Any patent application processing fees under 37 C.F.R. § 1.17.

A check to cover the filing fee is enclosed.

The Commissioner is hereby authorized to charge payment of the following fees during the pendency of this application or credit any overpayment to Deposit Account No. 12-2325. A duplicate copy of this sheet is enclosed.

Any patent application processing fees under 37 C.F.R. § 1.17.  
 The issue fee set in 37 C.F.R. § 1.18 at or before mailing of the Notice of Allowance, pursuant to 37 C.F.R. § 1.311(b).  
 Any filing fees under 37 C.F.R. § 1.16 for presentation of extra claims.

Respectfully Submitted,

Eunhee Park  
Reg. No. 42,976  
BAKER & MCKENZIE  
805 THIRD AVENUE  
NEW YORK, N.Y. 10022  
212-751-5700

Date: March 7, 2000

EXPRESS MAIL NO.: EE225750495US  
DATE OF DEPOSIT: March 7, 2000

**Application  
For  
United States Letters Patent**

**To all whom it may concern:**

Be it known that I,

**Garry Z. Gu**

have invented certain new and useful improvements in

**CREDIT-BASED ADAPTIVE FLOW CONTROL FOR MULTI-STAGE MULTI-DIMENSIONAL SWITCHING ARCHITECTURE**

of which the following is a full, clear and exact description:

Eunhee Park  
Reg. No. 42,976  
Baker & McKenzie  
805 Third Avenue  
New York, NY 10022

[NYC] 322197.3

EXPRESS MAIL LABEL NO.: EE225750495US  
DATE OF DEPOSIT: MARCH 7, 2000

CREDIT-BASED ADAPTIVE FLOW CONTROL FOR MULTI-STAGE MULTI-  
DIMENSIONAL SWITCHING ARCHITECTURE

5

DESCRIPTION

TECHNICAL FIELD OF THE INVENTION

10 The present invention relates generally to switch architecture, and more particularly to credit-based adaptive traffic flow control for a multi-stage multi-dimensional switch.

BACKGROUND OF THE INVENTION

15

Conventional multi-stage, multi-dimensional switching architecture typically includes three stages. The first and third stages are the buffering stages while the second stage acts as a bufferless crossbar switch node. Each 20 stage includes an array of switching element and each element includes either a queuing processor or an arbitration processor which makes decisions to route arriving data packets from any input port to any output port.

25 Because of its easy scalability and expandability, the multi-stage, multi-dimensional switching technique has been acceptable in many types of switched architecture including many data switch applications. On the other hand, disadvantages of such switching techniques include complex 30 algorithms required for employing the switch resulting in a substantial switching latency. As generally recognized by those skilled in the art, the routing algorithms required in the multi-stage, multi-dimensional switch architecture

[NYC] 322197.3

EXPRESS MAIL LABEL: EE22575495US  
DATE OF DEPOSIT: March 7, 2000

are complex. Further, such routing algorithms become exponentially more complex with increasing number of switching elements within an architecture. The more complex algorithms require more complex internal data structures in the bidding stage (first stage) and the granting stage (second stage), additional switch tag fields, more complex bidding process, and ultimately longer time for a grant/reject response to a bid to turn around. The increasing number of switching elements within the architecture also creates greater potential contention for any bidding effort. These disadvantages associated with the multi-stage, multi-dimensional switch architecture have prompted a need for additional research in the field.

One example of the bottlenecks in the complex routing algorithms is that data packets from many or all input ports of the first-stage may arrive at a specific input port of the second-stage at the same time, resulting in a severe contention consequence. An active bidding and granting procedure is considered as a primary solution for this problem. Current research activities mainly concentrate on techniques for improving the active bidding and granting process, and pushing the design to the limit of state-of-the-art fabrication technology. However, many timing constraints resulting in bidding and granting logic as well as interfaces between these two stages make active bidding and granting logic hard to implement, and make the grant response to the bid very hard to happen within a reasonable time period. This problem becomes even worse when more and more switch elements are involved.

Therefore, it is highly desirable to have a self-adjustable flow control technique for the multi-stage, multi-dimensional switching architecture which does not require any specific data fields from the upstream switch element, and of which the complexity is not related to the number of increasing switching elements within a switch architecture.

Switches providing interconnections between nodes on a network invariably include some type of a traffic flow control technique. The credit-based flow control technique is one such traffic flow control technique. The credit-based flow control technique currently available in the market and generally known to those skilled in the art are typically designed for flow control between two switch elements, referred to as hubs, at a network level on a one-to-one basis. The hub-to-hub, credit-based, flow control usually resolves congestion earlier as compared to the end-to-end flow control techniques. Although different versions of this technique have been adopted in many switch products, the credit-based flow control technique has not yet been applied inside a switching hub.

Filtering techniques are generally known and have been applied in various applications. For example, an adaptive filtering technique has been applied to radar antenna arrays. Additionally, one-dimensional applications of adaptive filtering may be found in the field of electronics and also in the field of data routing known as adaptive directory routing. Furthermore, multi-dimensional adaptive filtering is currently being applied to digital image processing applications. Generally, however, the adaptive filtering technique has not yet been applied to a data switch design. Moreover, multi-dimensional filtering has not been applied to any type of a network design.

#### SUMMARY OF THE INVENTION

The present invention combines credit-based flow control, multi-stage multi-dimensional switching, and adaptive filtering together and seamlessly applies them to a mesh of switching elements. More particularly, the present invention adopts a multi-dimensional adaptive filtering mechanism into a multi-stage, multi-dimensional,

switching architecture.

The adaptive filtering process in one embodiment of the present invention is accomplished by adaptively adjusting a "threshold" array or grant credit. When 5 employed, the adaptive filtering process of the present invention drives the statistical properties of the network to conform to nearly uniform probability distribution.

In accordance with the goals of the present invention, there is provided a novel multi-stage, multi-dimensional, 10 credit-based adaptive flow control method. In one embodiment, each crossconnect element of the second stage accumulates the statistics of data packet arrivals and departures on per-port based, and notifies the first stage with a single token bit whenever the second stage receives 15 a data packet from a specific output port, for example, output port #i, of the first-stage, and departs it. The value of an up/down counter for this specific output port is maintained until the data packet is departed from the second stage.

20 The first stage switch elements update the grant credits associated with the output port #i when notified by the token bit from the second stage, and then waits for next data packet arrival from external nodes. The grant credit updates generally apply to three integrator blocks 25 associated with three neighboring output ports of the first stage, subtracting one credit from the output port which has just sent out a data packet to the second stage, and simultaneously adding one credit to two of its next neighboring integrator blocks.

30 The first stage switch elements use the updated grant credit information to decide to which second stage switch element, i.e., crossconnect element, the next data packet should be sent. The advantage of doing so is that the first stage knows where to route an incoming packet prior 35 to its arrival so that the active bidding alternative,

i.e., the first stage sends a bid to the second stage to request an access grant, is no longer needed. This leads to relaxing many stringent design requirements, and therefore, larger network meshes can be built as a result.

5       Upon every arrival at the first stage, the credits with all the output ports in each first stage element may be scanned in the ascending order from 0 to N-1 to find a maximum. The data packet is then sent via the output element with the maximum credit to the second stage.

10       Another aspect of the present invention includes a simple credit feedback loop embedded into the data stream from the second stage to the third stage, and the third stage relays the credit token back to the first stage. Further, the first stage and third stage may be on the same 15 chip or the same board.

20       Yet another aspect of the present invention includes a plurality of interconnected digital logic blocks including second stage statistic blocks, first stage integrator blocks, and first stage threshold blocks. The second stage statistics blocks are arranged to maintain the real-time statistics of data packet arrivals and departures. The first stage integrator blocks receive logic values from the statistics blocks and update grant credits accordingly. The first stage threshold blocks look for an output port 25 with maximum grant credits and forward the packets via the output port to second stage where feedback tokens are then generated and sent back to the first stage.

30       Further features and advantages of the present invention as well as the structure and operation of various embodiments of the present invention are described in detail below with reference to the accompanying drawings. In the drawings, like reference numbers indicate identical or functionally similar elements.

The embodiments of the present invention will now be described, by way of example only, with reference to the accompanying drawings in which:

5       Figure 1 shows a three-stage multi-dimensional switch architecture to which the method of the present invention is applied;

10      Figure 2 is a diagram illustrating the structure of credit-based adaptive flow control logic inside each switch element in Stage-1 and inside each crossconnect element in Stage-2 in one embodiment of the present invention;

15      Figure 3 shows the upper boundary structure of the credit-based adaptive flow control logic in one embodiment of the present invention;

20      Figure 4 shows the bottom boundary structure of the credit-based adaptive flow control logic in one embodiment of the present invention;

25      Figure 5 is a diagram illustrating a detailed internal logic flow of a two-port first stage switch element in one embodiment of the present invention; and

30      Figure 6 is a diagram illustrating a detailed internal logic flow of a second stage crossconnect element in one embodiment of the present invention.

25

#### DETAILED DESCRIPTION OF THE INVENTION

35      Figure 1 shows the architecture of a three-stage multi-dimensional switch fabric 100 in one embodiment of the present invention. Briefly, a switching element is a basic building block of a switch and generally includes devices for providing inputs and outputs. As generally known to those skilled in the art of network switches, a switching fabric refers to an interconnection of switching elements. As shown in Figure 1, the switch architecture

100 of the present invention includes at least three stages. The first stage 102 includes an array of switch elements 0 to L-1 (108a, 108n). The third stage also includes an array of switch elements 0 to L-1 (112a, 112n).  
5 Each switch element 108a, 108n has M physical input ports (or queues) 114 and N physical output ports 120 which route the incoming data to the next stage. Each third stage switch element 112a, 112n has N physical input ports (or queues) 118 and M physical output ports 122 which route the  
10 incoming data to an external node. Preferably, N is greater than M to ensure non-blocking capability of each switch element. The second stage includes an array of cross-connect elements 0 to N-1 (110a, 110n). Each cross-connect element (110a, 110n) includes L physical port  
15 processors 116 each performing the space-domain switching. Space-domain switching is known to those skilled in the art of network switches and generally describes internal switch routing structure that allows connections between every input to every output.

20 In one embodiment of the present invention, the credit-based adaptive flow control is applied between first stage 102 and second stage 104 to eliminate the conventional active bid-and-grant process. The credit-based adaptive flow control of the present invention will  
25 now be described in greater detail with reference to Figure 2.

Figure 2 illustrates the credit-based adaptive flow control logic 200 in the first stage ("Stage-1") 202 and second stage ("Stage-2") 204 of the present invention. In  
30 describing the present invention, the term "input element" will be used synonymously with an input port processor (114 Figure 1) while the term "output element" used synonymously with an output port processor (120 Figure 1). For example, elements 0 to M-1 of first stage refer to the port  
35 processors 0 to M-1 (114 Figure 1) of switch elements 0 to

L-1 (108a to 108n Figure 1) in the first stage as shown in Figure 1. Elements 0 to L-1 of second stage refer to port processors 0 to L-1 (116 Figure 1) of cross-connect elements 0 to N-1 (110a to 110n Figure 1) in the second stage. Similarly, input elements 0 to N-1 (118 Figure 1) of third stage refer to the input port processors 0 to M-1 (122 Figure 1) of switch elements 0 to L-1 (112a to 112n Figure 1) in the third stage.

Referring back to Figure 2, the first stage 202 includes a plurality of first stage output elements, each of which plurality includes a threshold block 216 and an integrator block 206. The second stage 204 includes a plurality of second stage elements, each of which plurality includes a statistics block 208 corresponding to the threshold block 216/integrator block 206 in the first stage.  $Grant_i(0)$  shown at 230 for  $i=0, \dots, M-1$ , of the integrator block is the initial value of  $Grant_i(k)$  when time k is equal to 0. The value of  $Grant_i(0)$  230 depends upon the switch latency of the second stage 204 and traffic class of an arriving data packet. Each integrator block 206 of Stage-1 202 accumulates the grant tokens received from Stage-2 204 with the initial value  $grant_i(0)$  230 as a starting point. Each integrator block 206 such as i-th integrator takes as input the logic values or token bits representing grant tokens from the corresponding statistics block 208, i.e., i-th block Stage-2, as well as from its neighboring blocks #(i-1) and #(i+1).

As shown in Figure 2, each statistics block 208 of Stage-2 accumulates the number of arrivals and departures by incrementing the statistics counter 214 by 1 for each arrival and by decrementing the statistics counter by 1 for each departure. A statistics block may include a simple up/down counter 214 for incrementing a count when a packet arrives, and decrementing a count when the packet departs. The counter 214 may be coupled to a logic circuit 232 which

evaluates to "true" value when the counter 214 is incremented and then decremented for a particular packet arrival and departure. The "true" value triggers a token bit to be sent back to the first stage switch elements as shown.

In the present invention, the second stage is bufferless, i.e., the second stage receives arriving data packets and transmits them to the third stage without storing or queuing the data packets in the second stage.

10 The arrival and departure of data packets in the second stage are maintained by the second stage statistics block 208. As described above, the second stage statistics block 208 is arranged to accumulate the real-time statistics of data packet arrivals and departures. Every new arrival 224  
15 and departure 226, to and from i-th element of Stage-2 204 triggers sending out a feedback token or a token bit, shown at 228, back to the i-th output element of Stage-1 202. When the new arrival departs Stage-2, the counter 214 is decremented by 1 as shown at 226. When a feedback token is  
20 triggered as a result of a new arrival, the grant credit of the i-th output element of Stage-1 is decreased by 1 while the grant credits of the (i-1)-th output element and (i+1)-th output element in Stage-1 are increased by 1. This unique adaptive filtering process effecting the credit-  
25 based adaptive flow control in a multi-stage multi-dimensional switching architecture is shown in Figure 2 by the arrows at 218, 220, 222 leading back to i-th, (i-1)-th, and (i+1)-th output elements in Stage-1, respectively. The arrow at 228 illustrates a logic value "a" as an example of  
30 a token bit being transmitted back to the first stage switch elements.

With the above-described method, statistics block 208 (e.g., i-th element) updates the threshold block 216 via the integrator block 206 when data packets arrive and  
35 depart at Stage-2. The corresponding first stage

integrator block 206 (e.g., i-th output element) receives logic values or token bits from the statistic block 208 and updates grant credits. Briefly, grant credits represent the availability of a second stage element to receive an 5 incoming data packet for routing. At the same time, the first stage integrator 236 that is above the corresponding first stage integrator 206 and the first stage integrator 234 that is below the first stage integrator 206 receive logic values or token bits from the statistics block 208 10 also and update their grant credits accordingly. When the filtering technique of the present invention described above is employed, the grant credits in the "threshold" blocks of the present invention converge to a nearly uniform probability distribution, and therefore the traffic 15 between the first stage and second stage becomes balanced.

Thus, in one embodiment of the present invention, a nearly uniform distribution probability is used as a reference in this adaptive "threshold" filtering technique.

The operation of a "threshold" filtering of the 20 present invention will now be described in greater detail herein below with reference to equations (1) to (7). The i-th element among the plurality of processors in the switch architecture represents an element being considered, (i-1)-th element represents an element above the i-th 25 element, and (i+1)-th element represents an element below the i-th element in the switch architecture.

By using statistical information derived from the normal switching operation, the filtering process of the present invention estimates data flow and balances them in 30 the multi-stage multi-dimensional data switch architecture without requiring specific information from the previous stage.

When  $grant_i(k) = \text{maximum among } \{grant_i(k), i = 0, \dots, N-1\}$  and the packet is thus sent from the i-th output 35 element of the first stage to the i-th element of Stage-2,

then the first stage does the following updates upon receiving the token bit from the second stage:

$$\text{grant}_i(k+1) = \text{grant}_i(k)-1, \quad (1)$$

$$\text{grant}_{i-1}(k+1) = \text{grant}_{i-1}(k)+1, \quad (2)$$

$$5 \quad \text{grant}_{i+1}(k+1) = \text{grant}_{i+1}(k)+1, \quad (3)$$

$i=0, \dots, N-1$ .

The following two end conditions are maintained.

For Upper Boundary, when  $\text{grant}_0(k) = \text{maximum}$  among

10  $\{\text{grant}_i(k), i = 0, \dots, N-1\}$  and the packet is thus sent to the 0-th element of Stage-2, the first stage integrator blocks perform the following operations upon receiving the token bit from the second stage:

$$\text{grant}_0(k+1) = \text{grant}_0(k)-1, \quad (4)$$

$$15 \quad \text{grant}_1(k+1) = \text{grant}_1(k)+1. \quad (5)$$

For Bottom Boundary, when  $\text{grant}_{M-1}(k) = \text{maximum}$  among

$\{\text{grant}_i(k), i=0, \dots, N-1\}$  and the packet is thus sent to the  $(M-1)$ -th element of Stage-2, the first stage integrator

20 blocks perform the following operations upon receiving the token bit from the second stage:

$$\text{grant}_{M-1}(k+1) = \text{grant}_{M-1}(k)-1, \quad (6)$$

$$\text{grant}_{M-2}(k+1) = \text{grant}_{M-2}(k)+1. \quad (7)$$

25 In one embodiment of the present invention, the search for the maximum among  $\{\text{grant}_i(k), i=0, \dots, N-1\}$  at each switch element of the first stage is performed in the ascending order from 0 to  $N-1$ .

Figure 3 shows the upper boundary structure of the 30 credit-based adaptive flow control logic. The integrator block #0 302 receives as input the logic values from the corresponding statistics block #0 306 of Stage-2 and its only neighboring block #1. The upper boundary grant credit update is shown by equations (4) and (5).

35 Figure 4 shows the bottom boundary structure of the

credit-based adaptive flow control logic. The detailed structure of the switch is the same as shown in Figure 2 with the exception that the (N-1)-th integrator block 402 receives as input the logic values from the corresponding 5 statistics block #(N-1) 404 and its only neighboring block #(N-2). The bottom boundary grant credit update is shown by equations (6) and (7).

Figure 5 is a diagram 500 illustrating a detailed internal structure and logic flow of a two-port first stage 10 switch element. Each MxN port processors (114, 120 Figure 1) of first stage includes an input port processor 502a, an output port processor 504a, and an integrator 206 and threshold 216 blocks 506a. The detailed logic flow of these devices was described herein above with reference to 15 Figure 2. Generally, the data arriving at the input port processor 502a is routed to the second stage element via the output port processor 504a. In one embodiment, the data buffer 508, the routing matrix 510, and the routing control 512 is shared among the input and output port 20 processors 502a, 502b, 504a, 504b. As shown, each output port processor (120 Figure 1) includes integrator and threshold blocks 506a, 506b coupled to the output port processors 504a, 504b. The detailed structure of the integrator and threshold block is shown at 506. The logic 25 flow of the integrator and threshold block was described herein above with reference to Figure 2.

Figure 6 is a diagram 600 illustrating a detailed internal structure and logic flow of a second stage crossconnect element. Each L port processor (116 Figure 1) 30 of second stage includes an input port processor 602a, an output port processor 604a, and a statistics block 606a. The detailed logic flow of these devices was described herein above with reference to Figure 2. Generally, when a data packet arrives at the input port processor 602a, the 35 statistics block 606a increments an up/down counter 214 by

one. When the data packet is routed to the output port processor 604a or 604b via the routing control 610 and routing matrix 608, the statistics block 214 decrements its up/down counter by one. When the up/down counter 214 is 5 greater than zero, it signifies a data arrival at the input port processor 602a; when the data packet departs at the output port processor 604a, a token bit is transmitted to corresponding and neighboring integrator blocks 206 of first stage elements.

10 While the invention has been particularly shown and described with respect to an embodiment thereof, it will be understood by those skilled in the art that the foregoing and other changes in form and details may be made therein without departing from the spirit and scope of the 15 invention.

50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000

CLAIMS

What is claimed is:

5

1. A multi-stage, multi-dimensional, credit-based, adaptive flow control switch comprising:

10 a plurality of first stage port processors, each one of the plurality of first stage port processors having an integrator block for receiving a token bit and updating a grant credit in response to receiving the token bit; and

15 a plurality of second stage port processors connected to the plurality of first stage port processors for receiving data packets from the plurality of first stage port processors, each one of the plurality of second stage port processors having a statistics block coupled to a corresponding integrator block;

20 the statistics block further coupled to one or more neighboring integrator blocks for transmitting a token bit to the corresponding integrator block and the one or more neighboring integrator blocks, the statistics block transmitting the token bit in response to a second stage port processor associated with the statistics block receiving a data packet from one of the plurality of the first stage port processors.

25 2. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 1, wherein the corresponding integrator block includes an (i)th integrator block and the one or more neighboring integrator blocks include an (i-1)th integrator block that is located above the corresponding (i)th integrator block.

30 3. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 2, wherein

the one or more neighboring integrator blocks further include an (i+1)th integrator block that is located below the corresponding (i)th integrator block.

5       4. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 3, wherein the one or more neighboring integrator blocks increase associated grant credits in response to receiving a token bit from the statistic block.

10

5. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 4, wherein the corresponding integrator block decreases an associated grant credit in response to receiving a token bit from the 15 statistic block.

15

6. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 5, wherein the plurality of first stage port processors includes one 20 or more input ports and one or more output ports, wherein the arriving data packet is received at one of the one or more input ports and routed to an output port having a maximum number of grant credit.

25

7. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 1, wherein the switch further includes:

      a plurality of third stage port processors connected to the plurality of second stage port processors; and

30       the statistic block further includes a real-time counter for accumulating real-time statistics of data packet arrivals from one of the plurality of first stage port processors and data packet departures to one of the plurality of third stage port processors.

35

8. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 7, wherein the switch further includes a relay from each one of the third stage port processors to the plurality of first stage port processors, and the token bit is transmitted to the first stage port processors by one of second stage port processor and the third stage port processor.

9. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 8, wherein the plurality of first stage switch elements and the plurality of third stage switch elements are embedded on a single chip.

15 10. An adaptive filtering method for controlling traffic flow in a multi-stage, multi-dimensional, switched network, comprising:

20 notifying a plurality of first stage port processors in response to receiving a data packet from one of the plurality of first stage port processors; and

25 updating grant credits associated with said one of the plurality of first stage port processors and one or more neighboring first stage port processors in response to the notifying.

25

11. The adaptive filtering method of claim 10, wherein the notifying includes:

30 notifying said one of the plurality of first stage port processors; and

35 notifying said one or more neighboring first stage port processors.

12. The adaptive filtering method of claim 11, wherein the updating grant credits includes:

35 decrementing a grant credit associated with said one

of the plurality of first stage port processors; and  
incrementing one or more grant credits associated  
respectively with the one or more neighboring first stage  
port processors.

5

13. The adaptive filtering method of claim 12,  
further including:

determining a port processor among the plurality of  
first stage port processors for forwarding an incoming data  
10 packet by selecting an output port processor having a  
maximum grant credit.

14. The adaptive filtering method of claim 13,  
further including:

15 receiving the incoming data packet at the determined  
port processor; and  
routing the data packet to a corresponding second  
stage port processor.

20 15. The adaptive filtering method of claim 10,  
wherein the one or more neighboring first stage port  
processors include a first stage port processor coupled  
adjacent to said one of the plurality of first stage port  
processors.

25

16. The multi-stage, multi-dimensional, credit-based,  
adaptive flow control switch as claimed in claim 1, wherein  
the one or more neighboring integrator blocks change  
associated grant credits in response to receiving a token  
30 bit from the statistic block.

17. The multi-stage, multi-dimensional, credit-based,  
adaptive flow control switch as claimed in claim 1, wherein  
the corresponding integrator block change an associated  
35 grant credit in response to receiving a token bit from the

statistic block.

18. The multi-stage, multi-dimensional, credit-based, adaptive flow control switch as claimed in claim 1, wherein  
5 the statistics block departs the data packet.

CREDIT-BASED ADAPTIVE FLOW CONTROL FOR MULTI-STAGE MULTI-DIMENSIONAL SWITCHING ARCHITECTURE

ABSTRACT OF THE INVENTION

5

The novel credit-based adaptive flow control method and system for multi-stage, multi-dimensional, data switching architecture. The method and system of the present invention estimates and balances data flow in the 10 switch architecture by using statistical information derived from the normal switching operation without requiring specific information from the previous stage. The multi-stage multi-dimensional switching architecture of the present invention includes three stages. The first and 15 third stages are the buffering stages, and the second stage acts as a buffer-less crossbar switch node. Each stage includes an array of switching elements, and each element includes either a queuing processor or an arbitration processor arranged to progressively refine estimates of the 20 data traffic from switch elements in the first stage by the second stage. The second stage maintains the statistic of valid traffic arrival and departure, and feedback grant information bit to the first stage. The first stage treats every grant information as credit for each data departure 25 to the second stage, and sends the data if there is a grant credit available. The adaptive filtering process adjusts a threshold adaptively, rendering the statistical properties to resemble a chosen probability distribution.

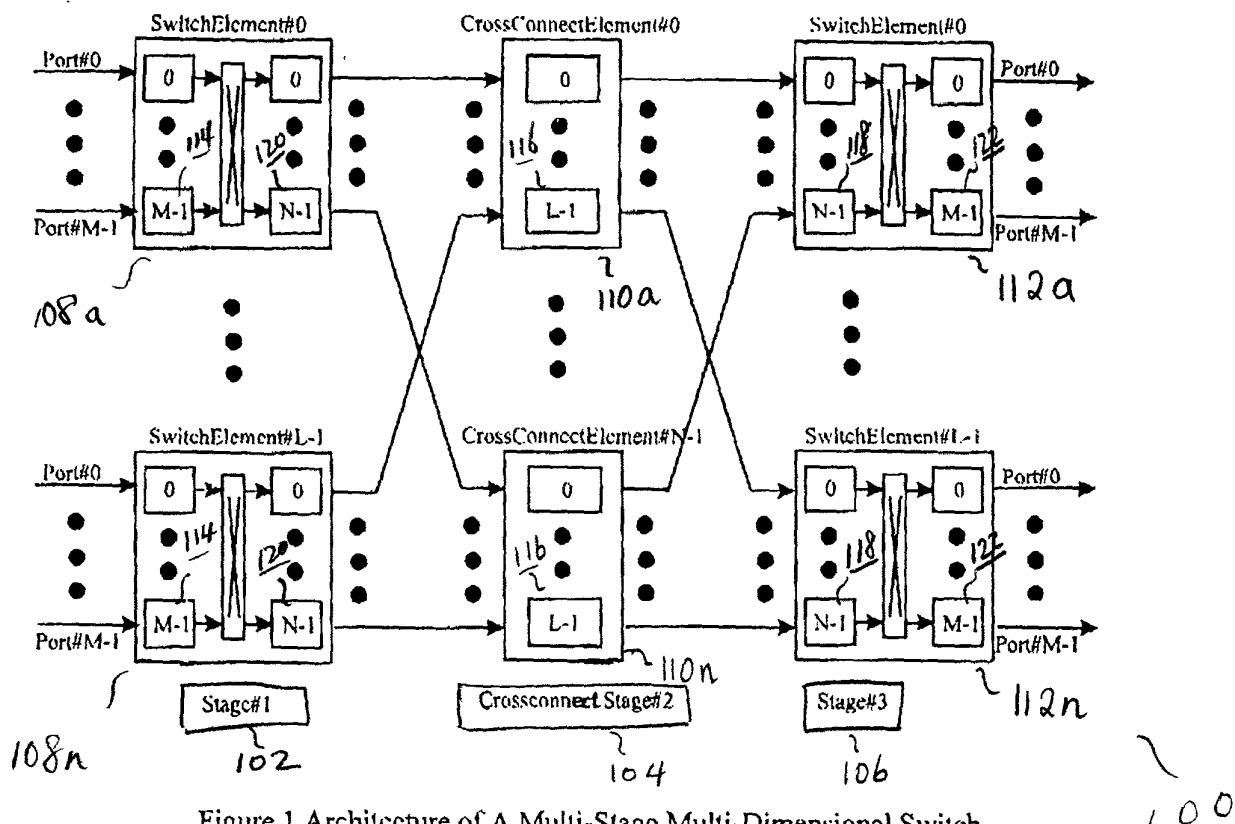


Figure 1 Architecture of A Multi-Stage Multi-Dimensional Switch  
to which the claimed technique is applied (usually  $N > M$ )

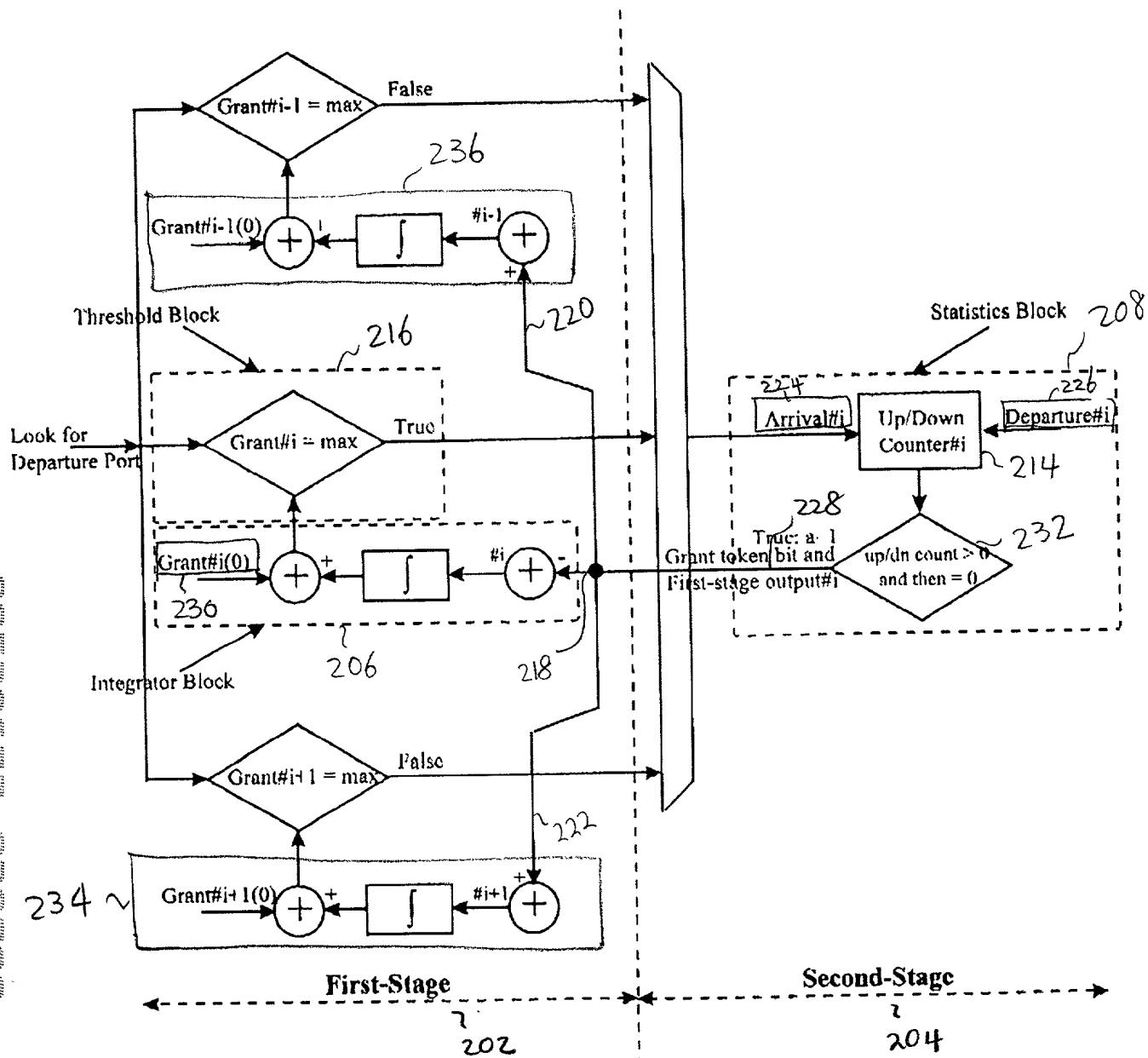


Figure 2 Structure of An Adaptive Credit-Based Flow Control Logic  
 Inside Each Switch Element at First Stage and Inside Each  
 Crossconnect Element at Second Stage When Data Packet  
 Arrives At Second-Stage From Input Port#i of First-Stage

300

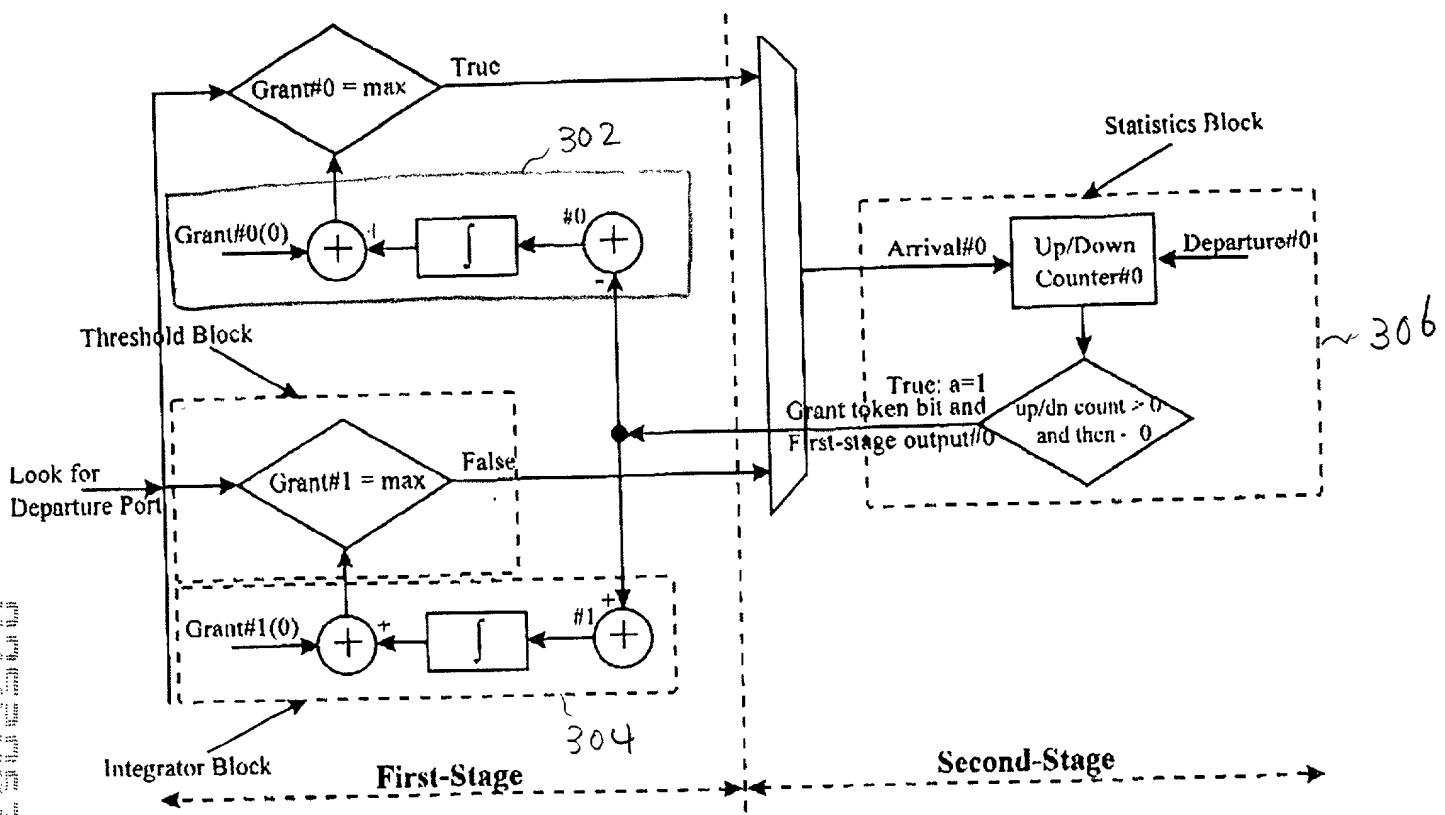


Figure 3 Upper Boundary of An Adaptive Credit-Based Flow Control Logic  
When Data Packet Arrives at Second-Stage From Output Port#0  
of First-Stage

400

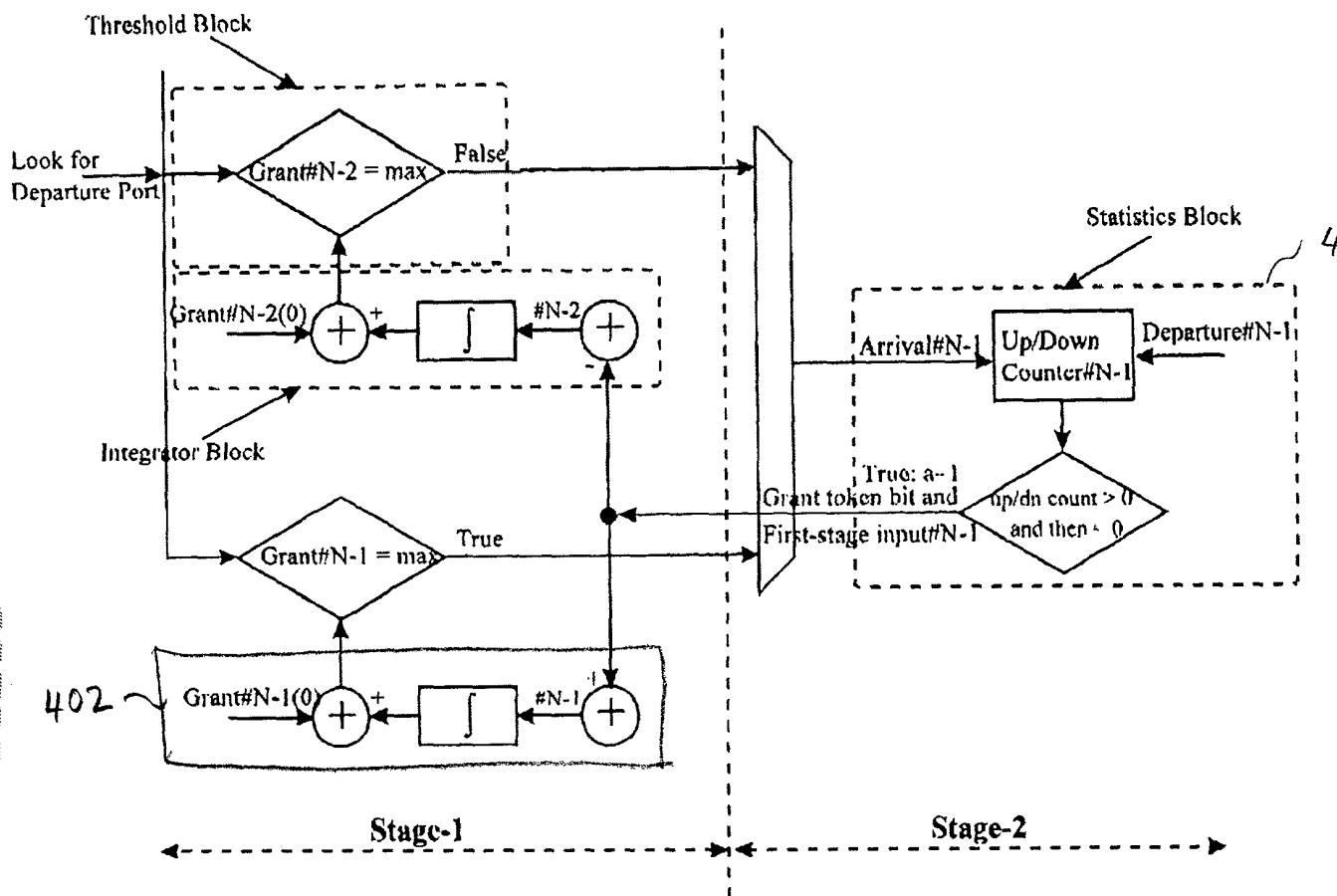


Figure 4 Bottom Boundary of An Adaptive Credit-Based Flow Control Logic  
When Data Packet Arrives at Second-Stage From Output Port#N-1  
of First-Stage

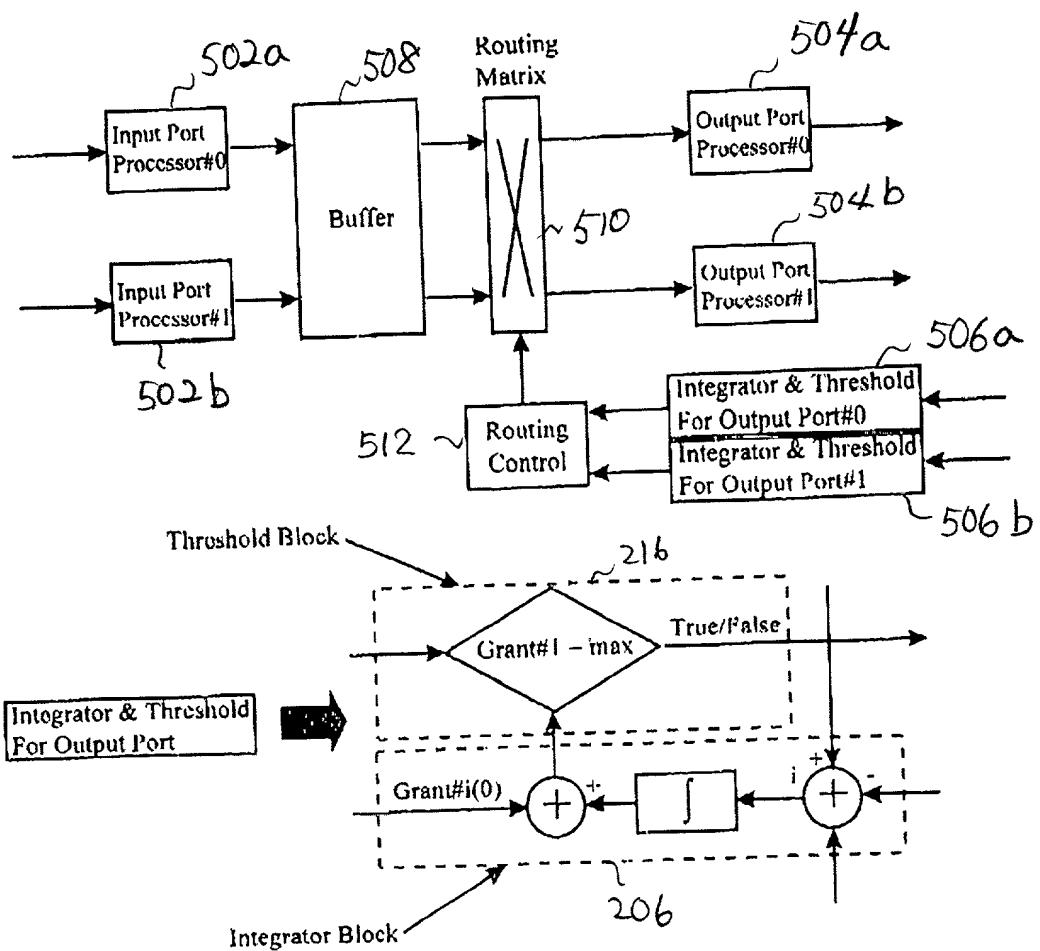


Figure 5 Internal Logic Flow of A Two-Port First-Stage Switch Element  
As Example

∫ 680

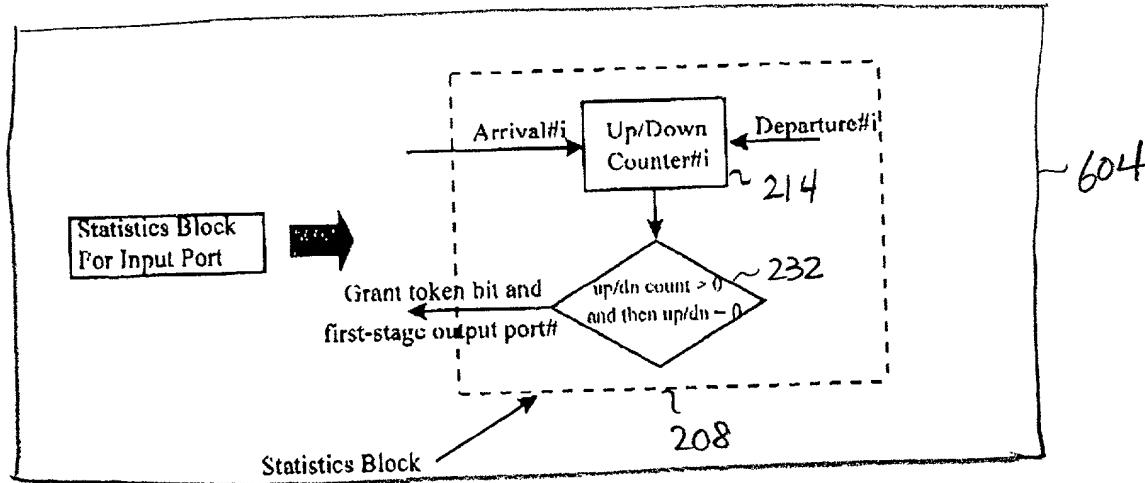
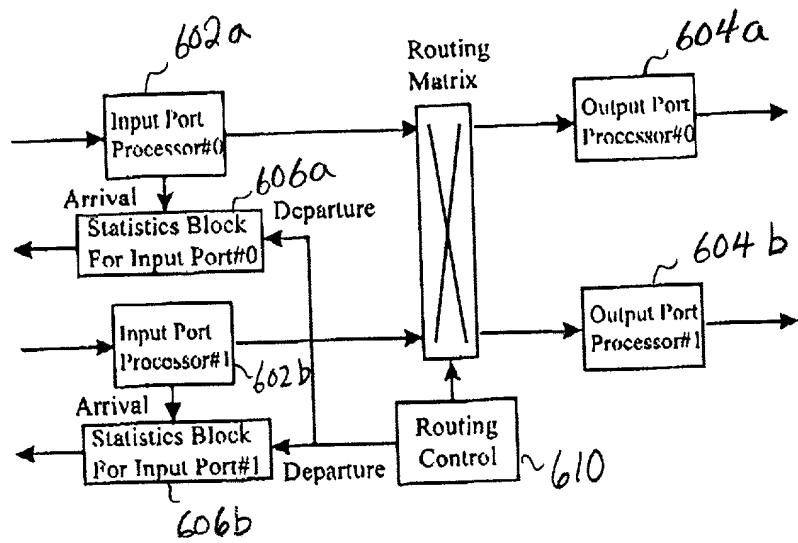


Figure 6 Internal Logic Flow of A Two-Port Second-Stage Crossconnect Element As Example

**DECLARATION AND POWER OF ATTORNEY**

As the below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe that I am the original, first inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled **CREDIT-BASED CONTROL FOR MULTI-STAGE MULTI-DIMENSIONAL SWITCHING ARCHITECTURE**, U.S. Patent Application No. filed herewith.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, § 1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application(s) for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

**PRIOR FOREIGN APPLICATION(S)**

Number	Country filed	Day/month/year	Priority Claimed Under 35 USC § 119
--------	------------------	----------------	--

I hereby claim the benefit under Title 35, United States Code, §§ 119(e) of any United States provisional application(s) listed below:

**PRIOR PROVISIONAL APPLICATION(S)**

Application Number	Filing Date
--------------------	-------------

I hereby claim the benefit under Title 35, United States Code § 120, of any United States application(s) listed below or under § 365(c) of any PCT international application(s) designating the United States of America listed below, and insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application(s) in the manner provided by the first paragraph of 35 U.S.C. 112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

**PARENT APPLICATION(S)**

U.S. Parent or PCT Parent  
Application Number

Parent Filing  
Date

Parent Patent Number  
(if applicable)

---

And I hereby appoint James David Jacobs (Reg. No. 24,299), Victor DeVito (Reg. No. 36,325), Harry K. Ahn (Reg. No. 40,243), Frank M. Gasparo (Reg. No. 44,700) and Eunhee Park (Reg. No. 42,976) as my attorneys with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

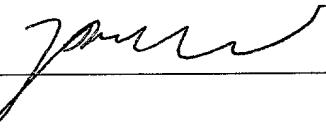
Please address all communications regarding this application to:

BAKER & MCKENZIE  
805 Third Avenue  
New York, New York 10022

Please direct all telephone calls to Victor DeVito at (212) 751-5700.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful and false statements may jeopardize the validity of the application or any patent issued thereon.

Sole Inventor: **Garry Z. Gu**

Inventor's Signature: 

Date: Feb 15, 2000

Residence: 4G-431  
101 Crawfords Corner Rd  
Holmdel, NJ 07733-3030

310 WEDGEWOOD ROAD  
MORGANVILLE, NJ 07751

Citizenship: USA CANADA

Post Office Address: Same as above.

  
**MARY L. BENSON**  
Notary Public of New Jersey  
My Commission Expires June 3, 2004